

Herbert H. Hyman

El modelo del experimento y el control de las variables

El experimento controlado como modelo

El modelo científico para el estudio de las relaciones de causa y efecto es el *experimento controlado*, en el cual se comparan las respuestas de un grupo experimental, expuesto al estímulo crucial, con las de un grupo de control equivalente, que no ha sido sometido al estímulo.

Esquemáticamente el experimento controlado puede ser representado del siguiente modo:

	Grupo experimental	Grupo de control
Momento 1	x_1	y_1
	Exposición al estímulo	Sin exposición al estímulo
Momento 2	x_2	y_2

Supongamos, por ejemplo, que nuestro experimento se ocupa del efecto que ejerce la televisión del desarrollo de las convenciones políticas sobre el interés por los asuntos políticos. En el Momento 1, antes de que comience la convención, dividiríamos nuestra muestra total en dos grupos equivalentes e igualados. Uno de éstos sería elegido para servir de grupo experimental y el otro de control. En el Momento 1 también registraríamos el nivel de interés político manifestado por los dos grupos en ese momento, y llamaríamos a esas dos medidas x_1 e y_1 . Si los dos grupos han sido igualados correctamente, deberían tener, en un comienzo, el mismo grado de interés; en otras palabras, x_1 sería igual a y_1 . A medida que las convenciones se desarrollan, se mostrará al grupo experimental las sesiones televisadas, mientras que, de algún modo, se impedirá que el grupo de control las vea. Una vez clausuradas las convenciones, se medirá

nuevamente el interés de ambos grupos con respecto a los asuntos políticos y se registrarán las nuevas mediciones x_2 y y_2 .

A partir de este esquema, ¿cómo se puede determinar si los programas de televisión han surtido efecto? Uno podría sentirse tentado a contestar que una comparación del nivel de interés manifestado por el grupo experimental antes y después de las convenciones proporcionará la información necesaria; en otras palabras, se podría decir que la comparación $(x_2 - x_1)$ es la crucial. Pero esto puede no ser del todo correcto. En el momento de las convenciones, los problemas políticos suelen ser de gran interés público. Otros medios de comunicación aparte de la televisión otorgan gran importancia a la marcha de las convenciones y a las especulaciones acerca de lo que ocurrirá después de finalizadas.

Los candidatos, sus posibilidades, sus méritos, se vuelven tópicos favoritos de conversación. Por lo tanto, a menos que se lo mantenga completamente aislado, el grupo experimental estará expuesto a estas influencias "extrañas", y la diferencia $(x_2 - x_1)$ reflejará tanto el efecto de estos factores como el del estímulo experimental. En otras palabras, es imposible distinguir qué parte de la diferencia $(x_2 - x_1)$ se debe a la televisación de las convenciones y qué parte a otros estímulos "irrelevantes".

Aquí es donde interviene el grupo de control. Presumiblemente ha estado expuesto a todos los mismos estímulos que el grupo experimental, con la única excepción de aquí en el que el experimentador está vitalmente interesado; en este caso, la televisión.¹ Así, la diferencia entre el nivel de interés exhibido por el grupo de control antes y después de las convenciones refleja la influencia de aquellos otros estímulos. Es decir, $(y_2 - y_1)$ mide el efecto de los estímulos no experimentales, aquellos que no son controlados por el experimentador. Estas dos diferencias, consideradas en conjunto, proporcionan una medida del efecto del estímulo experimentador verbalmente:

Efecto de la televisión = (efecto de la televisión + efecto de otros estímulos) - (efecto de otros estímulos)

¹ El experimentador debe suponer que el grupo de control ha sido expuesto a estos otros estímulos con la misma intensidad y con el mismo efecto que el grupo experimental. Si los dos grupos fueron bien igualados desde el principio, esta suposición probablemente es correcta.

o, simbólicamente:

$$\text{Efecto de la televisión} = (x_2 - x_1) - (y_2 - y_1)$$

Si los dos grupos fueron adecuadamente igualados al comienzo, de modo que inicialmente manifestaban su idéntico grado de interés político, esta medida del efecto se simplifica. Pues en estas condiciones, $x_1 = y_1$, y el efecto del estímulo se reduce a lo siguiente:

$$\text{Efecto de la televisión} = (x_2 - y_2)$$

La lógica de la experimentación controlada y los diseños experimentales específicos requeridos en diferentes condiciones han sido sistematizados ampliamente. Pero el estudiante verá que se ha prestado muy poca atención al estudio de los problemas que aparecen cuando se intenta la experimentación efectiva en ciencias sociales. El experimentador, por ejemplo, encuentra difícil manipular los grupos e instituciones sociales como querría para sus fines experimentales. Usualmente el experimentador de las ciencias sociales tampoco es capaz de mantener sobre sus sujetos los controles requeridos para una experimentación adecuada. El resultado es, entonces, que la experimentación se ha vuelto un ideal raramente realizado en las ciencias sociales. Continúa siendo el modelo con el que se juzga a otros diseños de investigación y en cuyos términos se evalúan los resultados de las investigaciones. Pero los estudios reales en las ciencias sociales generalmente se llevan a cabo empleando algún tipo de procedimiento aproximado.

Existen diversos modos de aproximarse a la lógica del experimento, pero el empleado más frecuentemente en el análisis de encuestas es el de comparaciones entre subgrupos. Esto implica una comparación de la frecuencia con que grupos caracterizados de diferentes maneras expresan una cierta actitud o exhiben una característica particular. Por ejemplo, en el estudio de Havemann y West sobre graduados del *college* se descubrió que los jóvenes que asistían a los así llamados *Ivy League Colleges** recibían en su carrera posterior salarios más altos que los graduados de otras instituciones. En tales comparaciones, el analista supone que los subgrupos que ha formado se aproximan

* Literalmente, Liga de la Hiedra, nombre con que se conoce a un grupo de instituciones de enseñanza superior de prestigio tradicional (Yale, Harvard, Columbia, Cornell, etc.). (N. del T.)

a los grupos experimental y de control de un experimento real, y que la característica que distingue a los diferentes grupos se aproxima al estímulo experimental.

La dificultad de equiparar la simple tabulación cruzada de un estudio con el experimento real reside en el peligro de que la relación obtenida sea espuria. A fin de introducir la noción de relación espuria, veamos cómo se podrían representar esquemáticamente estas comparaciones entre subgrupos.²

	Grupo experimental	Grupo de control
	Expuesto al estímulo	No expuesto al estímulo
Momento 2	x_2	y_2

Los "estímulos" en estas comparaciones son, para usar nuestro ejemplo anterior, tales como la clase de *college* en el que uno se graduó. Las actitudes o conductas críticas, registradas en las medidas x_2 e y_2 , son variables tales como ingreso anual. A pesar de que el "efecto" de estos estímulos podría ser evaluado de diversas maneras, la medida más simple y más apropiada para nuestros objetivos es la diferencia ($x_2 - y_2$).

Este diagrama, que es comparable al que representa la experimentación controlada, señala claramente algunas de las diferencias esenciales existentes entre la verdadera experimentación y el procedimiento aproximativo de comparación de subgrupos. El esquema sugiere que estas comparaciones constituyen una forma de experimentación trunca en la cual falta un paso decisivo. Recordemos que una de las características esenciales de la experimentación es que, a través de los procedimientos de igual-

² La representación esquemática de los diferentes procedimientos de aproximación sigue estrechamente los lineamientos de un artículo de Samuel A. Stouffer, titulado "Some observations on study design", *Amer. Journ. Sociol.*, I.V, 1949-50, pp. 355-361.

lación y azarificación, el experimentador se asegura de que sus dos grupos son inicialmente idénticos en todos los aspectos importantes.

El analista de encuesta, en cambio, se encuentra con sus grupos cuando éstos ya están constituidos (cuando ya han tenido la oportunidad o no de elegir sus puestos en el ejército, cuando ya se han graduado del *college*, o cuando han sido residentes de una comunidad de vivienda durante diversos períodos de tiempo). En otras palabras, no tiene la oportunidad de controlar de antemano la composición de sus grupos "experimentales" y de "control" de manera de estar seguro de que inicialmente son idénticos.³ En consecuencia, siempre existe el peligro de que las relaciones que el analista encuentra en los datos de su estudio sean espurias, y surjan de diferencias iniciales entre los grupos comparados. El así llamado "estímulo" puede aparecer erróneamente como causando un efecto que se debe en realidad a las diferencias iniciales entre los grupos. Dado que destruyen lo que de otro modo parecerían ser relaciones significativas, los factores que determinan las diferencias iniciales entre los subgrupos estudiados serán llamados "factores invalidantes".

Consideremos este problema más concretamente. En el estudio de West sobre los graduados del *college*, por ejemplo, inmediatamente sospechamos que la relación que encuentra entre el tipo de *college* al que se asiste e ingreso posterior es espuria. Podemos preguntarnos inmediatamente si los estudiantes de los *Ivy League Colleges* no eran más ricos desde el comienzo que la generalidad de los graduados de otros *colleges*. Los altos costos en las instituciones de la *Ivy League*, junto con las políticas de admisión de esos *colleges*, favorecen a los hijos de familias pudientes. Y estas familias pudientes pueden, a su vez, proporcionar a sus hijos los contactos sociales y profesionales a través de los cuales pueden alcanzar más fácilmente empleos mejor remunerados. En otras palabras, puede parecer que existe una relación entre

³ Esta afirmación no es, por supuesto, apropiada para los estudios explicativos que segulan el principio de la *restricción del universo* presentado en el capítulo II. Bajo tales condiciones, el universo del que provienen todos los subgrupos es relativamente homogéneo y por lo tanto los subgrupos están, por definición, igualados en ciertos aspectos principales. Sin embargo, en el capítulo II también analizamos los muchos casos en que este principio de un universo restringido no es o no puede ser tomado en cuenta, y la necesidad de intentar un control *post-factum*. El análisis de toda esta sección es apropiado a un estudio explicativo de este último tipo.

ser graduado en un *Ivy League College* y éxito financiero sólo porque el egresado de estos establecimientos proviene de una familia pudiente que, a su vez, lo ayuda a lograr una posición financiera próspera.

La encuesta, que desembocó en el paradójico hallazgo de que los alemanes que habían leído *Mein Kampf* manifestaban actitudes más liberales que los que no lo habían hecho, constituye otro ejemplo del problema de la espuriedad. Al discutir las posibles secuencias temporales implicadas en esta relación empírica, resultaba que cada una de las secuencias nos enfrentaba con un aparente absurdo. Pero ahora podemos ver que este hallazgo tenía sentido. Estudios sobre la comunicación han dejado bien sentado que la lectura *per se* es una pauta generalizada que depende firmemente de la educación formal.⁴ Consecuentemente, los individuos que habían leído *Mein Kampf* probablemente tenían una educación superior a la de aquellos que no lo habían leído —no porque exista ningún interés inherente por *Mein Kampf* entre los más educados, sino simplemente porque existe una tendencia inherente a leer más. La educación, o más bien la afición a los libros fue el factor no controlado en la comparación entre los lectores y no lectores de *Mein Kampf*. Dado que esta gente muy probablemente también manifiesta esa particular actitud liberal (se les había pedido que dijeran si estaban de acuerdo o en contra de la siguiente afirmación: "El Estado puede quitar al pueblo sus derechos individuales si es para bien de todos"), empíricamente resultó que el efecto aparente de *Mein Kampf* fue generar liberalismo.

Así, a fin de minimizar el peligro de que tales relaciones espurias permanezcan ocultas, empleamos procedimientos analíticos que nos permiten examinar la relación entre la causa supuesta y el efecto supuesto cuando la influencia del posible factor invalidante es eliminada. Esto es, tratamos de eliminar las diferencias iniciales entre los subgrupos que podrían producir relaciones espurias. Los procedimientos analíticos para lograr este objetivo implican alguna manera de "mantener constantes" o "controlar" los posibles factores invalidantes.

⁴ Ver, por ejemplo, B. Berelson, *The Library's Public*, Columbia University Press, 1949, caps. I y II.

Problemas y métodos de control en la encuesta

El suicidio de Durkheim proporciona muchos ejemplos del peligro de la espuriedad y encara de un modo claro y explícito la solución técnica del problema en la investigación por encuestas. También ilustra ciertos rasgos especiales nuevos del tratamiento de la espuriedad. La naturaleza de la investigación de Durkheim era tal que el peligro de espuriedad resultaba *acentuado*. Se trataba de un análisis secundario —los datos obtenidos por Durkheim eran registros de pasados suicidios, algunas veces tabulados de tal modo que podían examinar las variaciones en las tasas para diversos subgrupos, por ejemplo región, período de tiempo. Obviamente, estos subgrupos variaban en muchos aspectos, pero a menudo Durkheim no podía encontrar una serie estadística para formar subgrupos muy refinados que fueran semejantes en muchos aspectos. En consecuencia, hizo frente al problema mediante una variedad de aproximaciones ingeniosas.

El suicidio de Durkheim. Estudio sobre los problemas de control.

Por la claridad de su exposición, ilustraremos el problema con sus propias palabras:

Consultando sólo las cifras absolutas, las personas solteras parecen suicidarse menos que las casadas. Así, en Francia, durante el período 1873-78 hubo 16.264 suicidios de personas casadas, mientras que entre las solteras hubo sólo 11.709. El primer número es al segundo como 132 es a 100. Como la misma proporción aparece en los otros períodos y en otros países, ciertos autores han sostenido que el matrimonio y la vida familiar multiplican las probabilidades de suicidio. Ciertamente, si de acuerdo con la opinión corriente se considera al suicidio primariamente como un acto de desesperación causado por las dificultades de la existencia, esta opinión parece muy probable. Una persona soltera tiene, de hecho, una vida más fácil que una casada. ¿No acarrea el matrimonio toda clase de cargas y responsabilidades? Para asegurar el presente y el futuro de una familia, ¿no se requieren más privaciones y sufrimientos que al enfrentar las necesidades de una persona sola? Sin embargo, a pesar de lo evidente que

parece, este razonamiento a priori es falso y los hechos sólo parecen sostenerlo en la medida en que han sido insuficientemente analizados... Realmente, para apreciar las cifras citadas debemos recordar que gran número de las personas solteras tienen menos de 16 años, mientras que todas las personas casadas son mayores. Hasta los 16 años la tendencia al suicidio es muy débil debido a la edad, sin considerar otros factores. En Francia se encuentran sólo uno o dos suicidios por millón de habitantes en esta época de la vida; en el período siguiente hay veinte veces más. De este modo, la inclusión de muchos niños menores de 16 años entre las personas solteras reduce indebidamente la aptitud promedio de estas últimas, dado que la reducción se debe a la edad y no al celibato. Si parecen suicidarse menos, esto no se debe a su soltería, sino al hecho de que muchas de ellas aún son inmaduras. Así, *si se trata de comparar las dos poblaciones para determinar la influencia del status marital, y sólo de éste, uno debe librarse de este elemento perturbador y comparar las personas casadas sólo con las solteras de más de 16 años.* Cuando se realiza esta substracción, surge que entre 1863-68 hubo un promedio de 173 suicidios en un millón de personas solteras de más de 16 años, y 154,5 en un millón de personas casadas. La relación entre el primer número y el segundo es de 112 a 100.⁵

Pero Durkheim continúa y nos muestra que aunque la solución formal al problema de la espuriedad estaba bien fundada, existía una dificultad técnica aún no resuelta, que se aplica en general a todos los procedimientos de control o igualación a fin de eliminar la influencia de los factores invalidantes. Se trata simplemente del problema de la *tosquedad de las mediciones* que limita la precisión del control:

Hay así un cierto aumento debido al celibato. Pero es mucho mayor de lo que muestran las cifras precedentes. Realmente, hemos supuesto que todas las personas solteras de más de 16 años y todas las casadas tenían la misma edad promedio. Esto no es verdad. La mayoría de los hombres solteros en Francia, exactamente el 58 por ciento, están entre los 15 y 20 años, la mayoría de las mujeres solteras, exactamente el

⁵ Versión castellana, Schapire, Buenos Aires, 1965, p. 131. El subrayado es nuestro.

57 por ciento, tienen menos de 25 años. La edad promedio de todos los hombres solteros es 26,8 y la de todas las mujeres solteras 28,4. La edad promedio de las personas casadas, por el contrario, está entre los 40 y los 45 años. Para ambos sexos, el suicidio aumenta de acuerdo a la edad del siguiente modo:

De 16 a 21 años	45,9 suicidios por millón de habitantes
De 21 a 30 años	97,9 suicidios por millón de habitantes
De 31 a 40 años	114,5 suicidios por millón de habitantes
De 41 a 50 años	164,4 suicidios por millón de habitantes

Estas cifras se refieren a los años 1948-57. Si la edad fuera la única influencia, la aptitud para el suicidio de las personas solteras no podría estar por encima de 97,9 y la de las personas casadas estaría entre 114,5 y 164,4, es decir, alrededor de 140 suicidios por millón de habitantes. Los suicidios de las personas casadas serían a los de las solteras como 100 es a 69. Estos últimos serían sólo dos tercios de los primeros, pero sabemos bien que en realidad son más numerosos. Así, la vida familiar tiene por efecto invertir la relación. Sin el efecto de la vida familiar las personas casadas deberían suicidarse con una frecuencia 50 por ciento mayor que las solteras en virtud de su edad; sin embargo, el suicidio se da en ellas con una frecuencia apreciablemente menor. Así, podría decirse que el matrimonio reduce a la mitad el peligro del suicidio o, más exactamente, que la soltería produce un aumento expresado por la proporción 112/69, o 1,6. Así, si representamos la tendencia al suicidio de las personas casadas por la unidad, la de las personas solteras de la misma edad promedio debe ser estimada en 1,6.*

Como Durkheim lo destaca, las diferencias encontradas después de haber establecido controles todavía pueden representar la influencia del factor invalidante, si la igualación de los grupos ha sido tosca. La influencia de la edad ha sido reducida, pero no eliminada. En este caso particular la tosquedad del control, afortunadamente, no oscurece la relación general porque una vez establecido aquél el resultado inicial se *invierte*. A pesar del hecho de que el grupo de los solteros todavía es algo más joven y por lo tanto está *menos* predispuesto al suicidio en

* *Ibid.*, p. 132.

virtud del factor residual no controlado de la edad, la comparación revela ahora una tasa *mayor* de suicidio entre ellos. Así se obtiene un resultado definido y muy convincente porque queda demostrado a pesar del obstáculo que significa la tosquedad del control sobre la edad. Pero hay muchos casos en los que el problema de la tosquedad del control hace que la relación permanezca realmente indefinida. Tal caso ocurriría si después de aplicar el procedimiento de control la relación inicial se viera reducida en alguna medida.⁷ El analista se enfrentaría entonces con el siguiente dilema: ¿representará el nuevo resultado la explicación válida, o la diferencia podría todavía ser reducida y *finalmente* invertida si se establecieran controles cada vez más refinados? Si no existe ningún método empírico para establecer tal control adicional, y dado que no existe ningún método lógico para resolver el dilema, el resultado es en verdad indefinido. Afortunadamente, en análisis *primarios*, el investigador tiene la oportunidad de hacer una medición refinada de cualquier variable, y así puede teóricamente resolver el problema. En la práctica, sin embargo, las mediciones y consecuentemente los controles pueden ser toscos, y el ejemplo citado sirve para recordarnos un problema técnico general. Donde el investigador está limitado a un análisis *secundario*, sin embargo, el problema es serio. Aun teniendo las mejores intenciones, un grado ideal de control puede ser imposible de alcanzar. Durkheim destaca:

... con los datos disponibles, este método de cálculo es el único aplicable a la mayoría de los casos. ... Por consiguiente, si debe usarse para establecer la situación general, sus resultados no pueden ser sino aproximados. Naturalmente, es suficiente mostrar que la soltería aumenta la tendencia al suicidio; pero sólo da una idea muy inexacta de la magnitud de este aumento... El único modo de evitar estas dificultades es determinar la tasa para cada grupo por separado en cada edad. En tales condiciones se puede, por ejemplo,

⁷ Es fácil advertir que una pauta de resultados distinta sería definida. Si una vez establecido el control no se observara *ningún cambio* con respecto al resultado original, esto permitiría al analista sacar una conclusión a pesar de la tosquedad de la igualación. Si el factor controlado tuviera alguna importancia, su efecto debería quedar eliminado en parte al establecerse algún control, por tosco que éste sea. En consecuencia, el analista afirma que dicho factor no puede haber operado inicialmente.

comparar personas solteras de 25 a 30 años con personas casadas y viudas de la misma edad y hacer lo mismo para otros períodos; de este modo la influencia del status marital estará aislada de todas las otras influencias y aparecerán todas sus posibles variaciones. Por otra parte, este es el método que Bertillon aplicó primero a la mortalidad y a la tasa de nupcialidad. *Desgraciadamente, las publicaciones oficiales no contienen los datos necesarios para hacer esta comparación. En realidad, presentan la edad de los suicidas independientemente de su status marital.* Que nosotros sepamos, la única publicación que ha seguido una práctica diferente es la del gran ducado de Oldenburg (incluyendo los principados de Lubeck y Birkenfeld). Para los años 1871-85 esta publicación nos da la distribución de los suicidios por edad para cada categoría de status marital por separado. Pero este pequeño estado tuvo sólo 1.869 suicidios durante estos quince años. Como nada cierto puede concluirse de tan pocos casos, decidimos realizar el trabajo nosotros mismos, aplicado a Francia, con la ayuda de documentos inéditos en poder del Ministerio de Justicia. Analizamos los años 1889, 1890 y 1891. De este modo clasificamos alrededor de 25.000 suicidios.

Estos problemas del analista de fuentes secundarias hacen que, incidentalmente, Durkheim termine su pasaje con esta observación:

Sin embargo la tarea de reunir estos datos, considerable si la lleva a cabo un individuo solo, podría ser realizada por las oficinas oficiales de estadística. Se da toda clase de información sin valor y sólo se omite aquella que, como se verá más adelante, podría mostrar el estado de la vida familiar en las diferentes sociedades europeas.*

En este ejemplo de Durkheim, y en los ejemplos anteriores, los procedimientos *empíricos* de control se lograron igualando subgrupos en ciertos aspectos. El procedimiento de Durkheim, pese a que proporciona un control o igualación toscos, seguía sin embargo las líneas del análisis ortodoxo. La solución era simple y directa y Durkheim se limitó a seguir el modico prescripto.

* *Ibid.*, pp. 153-155. El subrayado es nuestro.

Pero antes dijimos que a causa de las limitaciones de sus fuentes, Durkheim tuvo que recurrir a otras aproximaciones para tratar la espuriedad. Examinar tales aproximaciones es muy provechoso para el analista ya que, cuando trabaja en análisis secundarios, la necesidad puede forzarlo a tomar la posición de Durkheim. Y aun cuando realice un análisis primario, puede considerar, por falta de previsión —lo que es muy común—, que le habría gustado controlar cierto factor en una comparación determinada, pero que, desgraciadamente, no disponía de los datos ideales para una igualación real de los grupos en una comparación refinada posterior. ¿Cómo ha de encarar entonces el analista el problema de la espuriedad en ausencia de procedimientos de igualación de subgrupos? La respuesta es simple en *principio*: reemplazando los métodos *empíricos directos* por varios *métodos indirectos*.

El tipo más frecuente de método indirecto todavía requiere procedimientos *empíricos* de igualación de grupos de control, pero *con respecto a una variable que se presume relacionada con la variable que el analista habría querido controlar*. El método es indirecto en el sentido de que se supone que la primera variable puede reemplazar a la segunda. Si esta suposición es justificada, el ulterior procedimiento de comparación entre subgrupos igualados es válido. Encontramos un ejemplo de este procedimiento indirecto en el estudio de Havemann y West.

West advertía que la relación que había encontrado entre tipo de *college* y sueldo anual era tal vez espuria debido a la existencia de diferencias en los orígenes familiares de los estudiantes comparados. En consecuencia intentó controlar ese factor, que podría haber invalidado la relación encontrada. No podía hacerlo de manera directa, porque la cédula usada en la encuesta no incluía ninguna medida directa del origen familiar del estudiante en la época en que éste iba al *college*. Sin embargo, tenía una medida indirecta: la necesidad del estudiante de trabajar para costearse sus estudios en parte o totalmente. West razonó de la siguiente manera: un estudiante que se ve obligado a trabajar para pagar todos o parte de sus gastos de educación provendrá, término medio, de una familia menos pudiente que un estudiante mantenido enteramente por sus padres durante sus estudios.

Por supuesto, éste no es un índice infalible de riqueza familiar. Algunas familias muy pobres pueden hacer grandes sacrificios para mantener a sus hijos en el *college*; en otros casos, algunas

familias pudientes pueden considerar que es "honroso" o un "buen entrenamiento" para sus hijos tener que ganarse parte de sus gastos del *college*.⁹

Si bien no debía suponer una correspondencia exacta entre el origen familiar y la necesidad de trabajar para pagar por lo menos parte de los gastos de educación, West pensó que esta última era un índice suficientemente bueno para sus fines. Lo que hizo entonces fue tomar dos grupos extremos del grupo total de hombres de 40 o más años: los que habían pagado con su trabajo *más de la mitad* de sus gastos de *college*, y los que no habían trabajado *en absoluto* para pagar sus gastos de educación. Dentro de cada uno de estos grupos la investigadora estudió entonces la relación entre el tipo de institución en que se había estudiado y los ingresos anuales.

Encontró así que la relación que había observado originalmente aún persistía cuando se mantenía constante el posible factor invalidante. Es decir, cuando pudo igualar a los subgrupos con respecto a la situación de sus familias anterior a su ingreso al *college* (medida por la necesidad de trabajar para pagar sus gastos de educación), todavía encontró que el tipo de *college* al que se asistía tenía un "efecto" sobre sus ingresos ulteriores. Pudo confiar así en que la relación original que había anunciado no estaba invalidada por ese factor.

⁹ Además, la variable "trabajar para ganarse ~~sus~~ propios gastos de educación" es una de aquellas variables ambiguas que pueden tomar distintos significados en diferentes contextos. En este caso es usada como una medida de la necesidad, y por lo tanto indirectamente como una medida de la situación familiar. También podría ser tomada como un índice de los tipos de experiencias adquiridas por el estudiante durante sus días de estudiante, *op. cit.*, cap. VI, cuadro 3.